

# PATENT ABSTRACTS OF JAPAN

(11)Publication number : **10-042264**

(43)Date of publication of application : **13.02.1998**

(51)Int.Cl.

H04N 7/15

G06F 3/14

H04M 3/56

(21)Application number : **08-193247**

(71)Applicant : **NEC CORP**

(22)Date of filing : **23.07.1996**

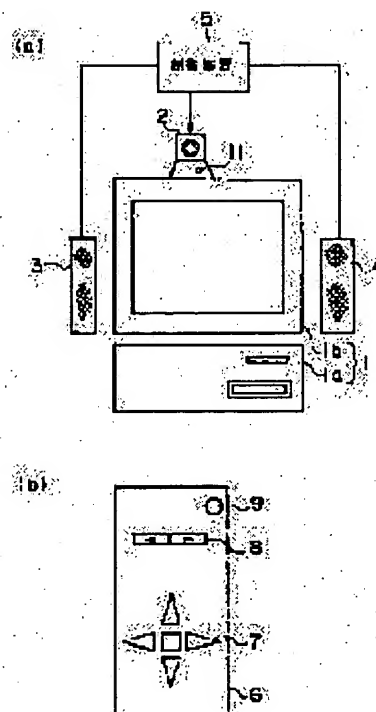
(72)Inventor : **NAKAMURA TERUYUKI**

## (54) VIDEO CONFERENCE SYSTEM

### (57)Abstract:

**PROBLEM TO BE SOLVED:** To eliminate the difficulty of recognizing who is speaking and the difficulty of performing an operation in a video conference system on a personal computer.

**SOLUTION:** Two microphones 3 and 4 serving also as speakers detect the position of the speaker and the television camera 2 of this video conference system is rotated in a detected direction. In which one of the left and right microphones detected voice is larger is compared first and the voice from which direction is the strongest in the microphone of a larger volume is detected. Then, the television camera is rotated in the detected direction. A conference participant can manually perform changeover by using a dedicated remote controller 6 when desired to fix or manually move the television camera during a conference.



## LEGAL STATUS

[Date of request for examination] 23.07.1996

[Date of sending the examiner's decision of rejection] 23.08.2000

[Kind of final disposal of application other than  
the examiner's decision of rejection or  
application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's 2000-15170  
decision of rejection]

[Date of requesting appeal against examiner's 22.09.2000  
decision of rejection]

[Date of extinction of right]

Copyright (C); 1998,2003 Japan Patent Office

\* NOTICES \*

JPO and NCIPi are not responsible for any damages caused by the use of this translation.

1. This document has been translated by computer. So the translation may not reflect the original precisely.
2. \*\*\*\* shows the word which can not be translated.
3. In the drawings, any words are not translated.

---

CLAIMS

---

[Claim(s)]

[Claim 1] In the video conference system containing the television camera connected with the personal computer arranged on the conference room at this A level difference detection means to detect whether a difference is in the receiver level of two sets of two sets of the microphones which set spacing and have been arranged in said conference room, and said microphones, A bearing specification means to pinpoint main speaker's of two or more attendants in said conference room existence bearing with reference to the difference when a difference is in said receiver level, The video conference system characterized by having a zoom out means to make said television camera start zoom out in the center of said conference room when there is no difference in said receiver level, the camera sense control means which doubles the sense of said television camera with said existence bearing, and.

[Claim 2] Each of two sets of said microphones is a video conference system according to claim 1 which is the microphone of loudspeaker combination.

[Claim 3] The video conference system according to claim 1 or 2 which carried two sets of said television camera and said microphones in the common frame.

[Claim 4] A video conference system given in either of claims 1-3 equipped with a zoom-in means to make said television camera cause zoom-in according to actuation of said camera sense control means.

[Claim 5] Said bearing specification means is a video conference system given in either of claims 1-4 which are what the receiver level of the corresponding microphone detects bearing which shows maximum, and makes said existence bearing after said main speaker judges in the location corresponding to any of two sets of said microphones it exists according to the difference of said receiver level.

[Claim 6] In the video conference system realized on the personal computer Two sets of the loudspeaker combination microphones which detect from which direction the voice at two or more attendants in the same conference room was uttered The video conference system characterized by being constituted with the camera which begins zoom out when it can rotate in the detected direction, a speaker can be tracked automatically and a microphone detects the sound volume of a constant rate from two or more places.

[Claim 7] The video conference system according to claim 6 constituted so that rotation to the four directions of said camera, halt, zoom-in, and actuation of zoom out could be performed as a remote controller is also.

---

[Translation done.]

\* NOTICES \*

JPO and NCIPI are not responsible for any damages caused by the use of this translation.

1. This document has been translated by computer. So the translation may not reflect the original precisely.
2. \*\*\*\* shows the word which can not be translated.
3. In the drawings, any words are not translated.

---

DETAILED DESCRIPTION

---

[Detailed Description of the Invention]

[0001]

[Field of the Invention] This invention relates to the video conference system which used the PASONARU computer.

[0002]

[Description of the Prior Art] When two or more attendants hold a conference as a conventional video conference system in the same conference room, there are some which introduce the focus of a camera to a speaker because one of the attendants of a board operates a remote controller.

[0003] Moreover, the microphone for board attendance is prepared, and those microphones are made to correspond to an attendant by one to one, it arranges, and the video conference system of an automatic-tracking mold which rotates a camera in the direction of the microphone with which voice was uttered is also proposed (refer to JP,5-122689,A or JP,61-198891,A).

[0004]

[Problem(s) to be Solved by the Invention] However, in the former, it has the problem that a camera lacks a speaker in presence rather than accuracy and the board which there is a limitation in following quickly and is held by meeting.

[0005] Moreover, the number of microphones increases in the latter and it is unsuitable at the television conference using a personal computer.

[0006] So, the technical problem of this invention is to offer the video conference system which photoed all board attendants, when a television camera tracks a main speaker automatically and a main speaker moreover cannot be specified, though it is the activity of a small number of microphone.

[0007]

[Means for Solving the Problem] In the video conference system containing the television camera which was connected with the personal computer arranged on the conference room at this according to this invention A level difference detection means to detect whether a difference is in the receiver level of two sets of two sets of the microphones which set spacing and have been arranged in said conference room, and said microphones, A bearing specification means to pinpoint main speaker's of two or more attendants in said conference room existence bearing with reference to the difference when a difference is in said receiver level, The video conference system characterized by having the camera sense control means which doubles the sense of said television camera with said existence bearing, and a zoom out means to make said television camera start zoom out in the center of said conference room when there is no difference in said receiver level is acquired.

[0008] Each of two sets of said microphones is good in it being the microphone of loudspeaker combination.

[0009] Two sets of said television camera and said microphones may be carried in a common frame. | ✕

[0010] It is desirable to have a zoom-in means to make said television camera cause zoom-in according to actuation of said camera sense control means.

[0011] Said bearing specification means of it being what the receiver level of the corresponding

microphone detects bearing which shows maximum, and makes said existence bearing after said main speaker judges in the location corresponding to any of two sets of said microphones it exists according to the difference of said receiver level is desirable.

[0012] Moreover, according to this invention, it sets to the video conference system realized on the personal computer. Two sets of the loudspeaker combination microphones which detect from which direction the voice at two or more attendants in the same conference room was uttered. It can rotate in the detected direction and a speaker can be tracked automatically, and when a microphone detects the sound volume of a constant rate from two or more places, the video conference system characterized by being constituted with the camera which begins zoom out is acquired. *rotatable*

[0013] in that case, rotation to the four directions of said camera, halt, zoom-in, and actuation of zoom out can be performed as a remote controller is also -- as -- configuration \*\*\*\* -- things are desirable.

[0014]

[Function] In the video conference system by this invention, while two sets of the loudspeaker combination microphones which set spacing and were installed hold a conference, a speaker's voice is detected. First, among microphones on either side, with reference to whether a speaker's voice is detected greatly, a speaker's bearing is detected to which microphone, and the rotation direction of a television camera is determined as. A television camera rotates in the determined direction, and photos a speaker. In this way, since, as for the attendant of remote \*\* of a television conference, it becomes easy for a speaker to judge someone by tracking the speaker under board automatically, presence which is holding a conference by meeting can be tasted.

[0015]

[Embodiment of the Invention] Drawing 1 is drawing for explaining the whole video conference system configuration by the gestalt of operation of this invention. This video conference system contains two sets of a personal computer (it abbreviates to a "personal computer" below) 1, a television camera 2, and the loudspeaker combination microphones 3 and 4. In order to hold a conference, it connects with a personal computer with the same system through an ISDN circuit etc.

[0016] A personal computer 1 contains body of computer 1a, and display unit 1b installed in the top face. A television camera 2 is installed in the top face of display unit 1b, and the neck swing of it is made possible under the left upper right respectively rotatable around the horizontal axis extended to vertical axes and right and left. Two sets of the loudspeaker combination microphones 3 and 4 are arranged at right and left of display unit 1b or a television camera 2, respectively. A television camera 2 and microphones 3 and 4 are connected to the control unit 5 by each. In addition, about actuation of a control unit 5, it mentions later. *rotatable*

[0017] The position of a television camera 2 and the sound volume of the loudspeaker combination microphones 3 and 4 are remotely controllable using a remote controller (it abbreviates to "remote control" below) 6. In this remote control 5, a sound-volume control switch for an attitude control switch for 7 to control the position of a television camera 2 and 8 to control the sound volume of the loudspeaker combination microphones 3 and 4 and 9 are the circuit changing switches for choosing whether control of a television camera 2 is performed automatically or it carries out manually.

[0018] Next, actuation of this video conference system is explained also with reference to the flow chart of drawing 2.

[0019] A control device 5 distinguishes first whether the circuit changing switch 9 shows automatic or hand control is shown at step S1. When the circuit changing switch 9 shows hand control, it will move to step S2 and remote control 6 will perform actuation of the neck swing under the left upper right of a television camera 2, zoom-in, and zoom out. In wishing to perform control of a television camera 2 automatically, a circuit changing switch 9 is operated and it chooses automatic. In addition, when a circuit changing switch 9 shows automatic, the automatic display lamp 11 of a television camera 2 shall come on red.

[0020] When the circuit changing switch 9 shows hand control at step S1, it moves to step S3 and judges whether a control device 5 has a difference in the receiver level of two sets of microphones 3 and 4. That is, an utterance of one board attendant judges with which microphone of the microphones 3 and 4 voice

is transmitted greatly. At this time, a control device 5 works as a level difference detection means.

[0021] When there is no difference in the receiver level of microphones 3 and 4, it moves to step S4, and a control device 5 turns a television camera in the center, and photos a lifting and all the members at the attendants of the board interior of a room by this for zoom out. At this time, a control unit 5 works as a zoom out means.

[0022] When a control device 5 detects a difference on the receiver level of two sets of microphones 3 and 4 at step S3, it moves to step S5 and detects whether it exists in the side to which a main speaker corresponds to which of two sets of microphones 3 and 4 among the attendants in a conference room based on the difference of the receiver level. Namely, since the man's voice will get across to a microphone 3 greatly compared with a microphone 4 supposing there is a main speaker on the left of the middle of a conference room, based on this, as for a control unit 5, a main speaker's existence location detects to which of two sets of microphones 3 and 4 it corresponds. Moving to step S6 furthermore, a control device 5 detects bearing where the receiver level of the corresponding microphone shows maximum, and pinpoints existence bearing. That is, bearing in which analyzes the source of utterance and a main speaker is present is pinpointed more in accuracy. When performing steps S5 and S6, a control unit 5 works as a bearing specification means. *detected difference*

[0023] Moving to step S7 furthermore, a control device 5 controls the position of a television camera 2 to suit pinpointed existence bearing, and photos a main speaker with the television camera 2. At this time, a control unit 5 works as a camera sense control means.

[0024] Finally, at step S8, a control device 5 causes zoom-in to a television camera 2, and performs photography. At this time, a control unit 5 works as a zoom-in means.

[0025] When another attendant speaks, it is the same approach as \*\*\*\*, and a television camera 2 rotates. That is, a television camera 3 always follows a main speaker.

[0026] By the way, it is in a camera 3 to track automatically. For example, it is the case where an attendant wants to explain data at hand. In such a case, automatic tracking is canceled by operating the circuit changing switch 9 of remote control 6, or clicking on the icon on display 1b of a personal computer 1 etc. Consequently, the position of a television camera 3 is manually controllable by operating the attitude control switch 7 of remote control 6.

[0027] Next, an example is given and it explains to a detail further.

[0028] The personal computer video conference system installed in each point through the ISDN circuit in two points is connected. 4-5 board attendants shall be in each point. *ISDN*

[0029] Attendant a1 who the board was started in the above examples and was sitting on the leftmost in the conference room of Point A. Supposing it speaks, he is an attendant a1 in the direction of which microphone between two sets of the loudspeaker combination microphones 3 and 4 of Point A. It compares whether voice is caught greatly. Since the left-hand side microphone 3 has caught greatly, he is the direction of left microphone 3 throat to the attendant a1 next. Voice is uttered, or the direction where the level of voice is the highest is investigated, and bearing is determined. According to this, a television camera 2 rotates, zooms in suitably and is photoed. Therefore, in Point B, he is an attendant a1. A face can be seen in the pictures and a main speaker understands immediately. *compare*

[0030] Now, nobody had the speaker at Point B at this time. In this case, since they can detect [ no ] voice from a direction, microphones 3 and 4 are these level, and a television camera 2 is turned in the center and they carry out zoom out of the microphones 3 and 4 on either side. That is, it becomes the form which photos the whole conference room.

[0031] Furthermore, another attendant a2 at Point A. Attendant a1 In spite of having talked, it began to talk. However, attendant a2 Voice is an attendant a1. Since it was small, it is not judged with this level, but a television camera 2 is an attendant a1 as it is. A photograph was taken.

[0032] Now, attendant b1 who thought below that he wanted to project a certain document on a document at Point B. The attendant who says began to talk. The focus of a television camera 2 is an attendant b1. Although it is suitable, he is an attendant b1. Since I wanted you to see the document, automatic-tracking mode was canceled and it was made to fix to the condition of having operated the television camera 2 with remote control, and having photoed the document. Thus, the board is advanced *cancel*

and automatic and hand control are changed.

[0033] attendant a1 of Point A or [ that two sets of the loudspeaker combination microphones 3 and 4 have gathered voice greatly with one of microphones first if it utters ] -- comparing -- a speaker -- right and left -- it judges in which side it is.

[0034] Next, comparison detection of from which direction a speaker's voice is uttered by the side in which the microphone has one microphone (microphone judged as a speaker being in near) is carried out, and the direction where audio level is the highest is specified as bearing where a main speaker exists.

[0035] Therefore, a television camera 2 will turn to the pinpointed bearing, and will photo a speaker.

*frame* [0036] Drawing 3 shows as a modification what built the loudspeaker combination microphones 3 and 4 and a television camera 2 into the frame 12 of one. And a voice tailing function is given to this microphone part by carrying a frame 12 in a truck 13 and making it movable. Moreover, it is desirable to enable it to rotate a television camera 2 360 degrees by making a frame 12 rotatable around vertical axes to a truck 13.

[0037] According to this, when it is made a manual mode, this equipment of one apparatus can be operated all around with remote control so that a camera can move data at hand.

[0038]

[Effect of the Invention] As explained above, according to this invention, the board without the presence in a personal computer has presence more. It is because the focus of a camera can be quickly introduced to a speaker at accuracy using two sets of loudspeaker combination microphones.

[0039] Moreover, control of a camera becomes easier. The reason is that it does not react to a speaker's voice that it became unnecessary to operate it so that a speaker may be settled in a screen, and by changing to hand control fixing the location of a camera by moving a camera automatically usually.

---

[Translation done.]

\* NOTICES \*

JPO and NCIPi are not responsible for any damages caused by the use of this translation.

- 1.This document has been translated by computer. So the translation may not reflect the original precisely.
- 2.\*\*\*\* shows the word which can not be translated.
- 3.In the drawings, any words are not translated.

---

TECHNICAL FIELD

---

[Field of the Invention] This invention relates to the video conference system which used the PASONARU computer.

---

[Translation done.]

\* NOTICES \*

JPO and NCIPi are not responsible for any damages caused by the use of this translation.

1. This document has been translated by computer. So the translation may not reflect the original precisely.
2. \*\*\*\* shows the word which can not be translated.
3. In the drawings, any words are not translated.

---

PRIOR ART

---

[Description of the Prior Art] When two or more attendants hold a conference as a conventional video conference system in the same conference room, there are some which introduce the focus of a camera to a speaker because one of the attendants of a board operates a remote controller.

[0003] Moreover, the microphone for board attendance is prepared, and those microphones are made to correspond to an attendant by one to one, it arranges, and the video conference system of an automatic-tracking mold which rotates a camera in the direction of the microphone with which voice was uttered is also proposed (refer to JP,5-122689,A or JP,61-198891,A).

---

[Translation done.]

\* NOTICES \*

JPO and NCIPi are not responsible for any damages caused by the use of this translation.

- 1.This document has been translated by computer. So the translation may not reflect the original precisely.
- 2.\*\*\*\* shows the word which can not be translated.
- 3.In the drawings, any words are not translated.

---

EFFECT OF THE INVENTION

---

[Effect of the Invention] As explained above, according to this invention, the board without the presence in a personal computer has presence more. It is because the focus of a camera can be quickly introduced to a speaker at accuracy using two sets of loudspeaker combination microphones.

[0039] Moreover, control of a camera becomes easier. The reason is that it does not react to a speaker's voice that it became unnecessary to operate it so that a speaker may be settled in a screen, and by changing to hand control fixing the location of a camera by moving a camera automatically usually.

---

[Translation done.]

JAPANESE

[JP,10-042264,A]

---

CLAIMS DETAILED DESCRIPTION TECHNICAL FIELD PRIOR ART EFFECT OF THE  
INVENTION TECHNICAL PROBLEM MEANS OPERATION DESCRIPTION OF DRAWINGS  
DRAWINGS

---

[Translation done.]

\* NOTICES \*

JPO and NCIPi are not responsible for any damages caused by the use of this translation.

- 1.This document has been translated by computer. So the translation may not reflect the original precisely.
- 2.\*\*\*\* shows the word which can not be translated.
- 3.In the drawings, any words are not translated.

---

MEANS

---

[Means for Solving the Problem] In the video conference system containing the television camera which was connected with the personal computer arranged on the conference room at this according to this invention A level difference detection means to detect whether a difference is in the receiver level of two sets of two sets of the microphones which set spacing and have been arranged in said conference room, and said microphones, A bearing specification means to pinpoint main speaker's of two or more attendants in said conference room existence bearing with reference to the difference when a difference is in said receiver level, The video conference system characterized by having the camera sense control means which doubles the sense of said television camera with said existence bearing, and a zoom out means to make said television camera start zoom out in the center of said conference room when there is no difference in said receiver level is acquired.

[0008] Each of two sets of said microphones is good in it being the microphone of loudspeaker combination.

[0009] Two sets of said television camera and said microphones may be carried in a common frame.

[0010] It is desirable to have a zoom-in means to make said television camera cause zoom-in according to actuation of said camera sense control means.

[0011] Said bearing specification means of it being what the receiver level of the corresponding microphone detects bearing which shows maximum, and makes said existence bearing after said main speaker judges in the location corresponding to any of two sets of said microphones it exists according to the difference of said receiver level is desirable.

[0012] Moreover, according to this invention, it sets to the video conference system realized on the personal computer. Two sets of the loudspeaker combination microphones which detect from which direction the voice at two or more attendants in the same conference room was uttered It can rotate in the detected direction and a speaker can be tracked automatically, and when a microphone detects the sound volume of a constant rate from two or more places, the video conference system characterized by being constituted with the camera which begins zoom out is acquired.

[0013] in that case, rotation to the four directions of said camera, halt, zoom-in, and actuation of zoom out can be performed as a remote controller is also -- as -- configuration \*\*\*\* -- things are desirable.

---

[Translation done.]

\* NOTICES \*

JPO and NCIPi are not responsible for any damages caused by the use of this translation.

1. This document has been translated by computer. So the translation may not reflect the original precisely.
2. \*\*\*\* shows the word which can not be translated.
3. In the drawings, any words are not translated.

---

DESCRIPTION OF DRAWINGS

---

[Brief Description of the Drawings]

[Drawing 1] It is the remote controller by which the configuration of the video conference system by the gestalt of operation of this invention is shown, (a) is used for the front view of the body and (b) is used for it.

[Drawing 2] It is a flow chart for explaining actuation of the control device in the video conference system of drawing 1.

[Drawing 3] It is the perspective view of only the important section for explaining the modification of this invention.

[Description of Notations]

1 Personal Computer

1a The body of a computer

1b Display unit

2 Television Camera

3 Microphone

4 Microphone

5 Control Unit

6 Remote Controller

7 Attitude Control Switch

8 Sound-Volume Control Switch

9 Circuit Changing Switch

11 Automatic Display Lamp

12 Frame

13 Truck

---

[Translation done.]

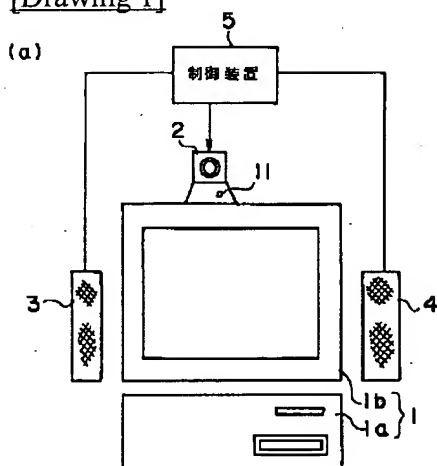
## \* NOTICES \*

JPO and NCIPi are not responsible for any damages caused by the use of this translation.

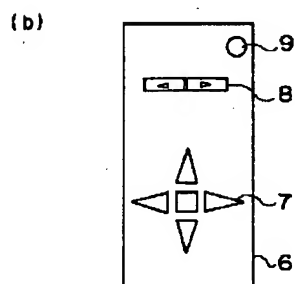
1. This document has been translated by computer. So the translation may not reflect the original precisely.
2. \*\*\*\* shows the word which can not be translated.
3. In the drawings, any words are not translated.

## DRAWINGS

[Drawing 1]

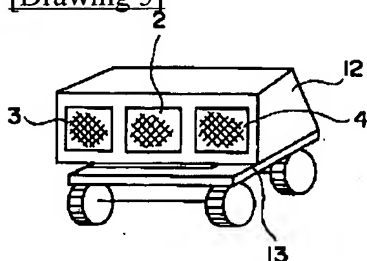


1: PC  
 2: TV Camera  
 3/4: Loud speaker & microphone  
 5: Controller

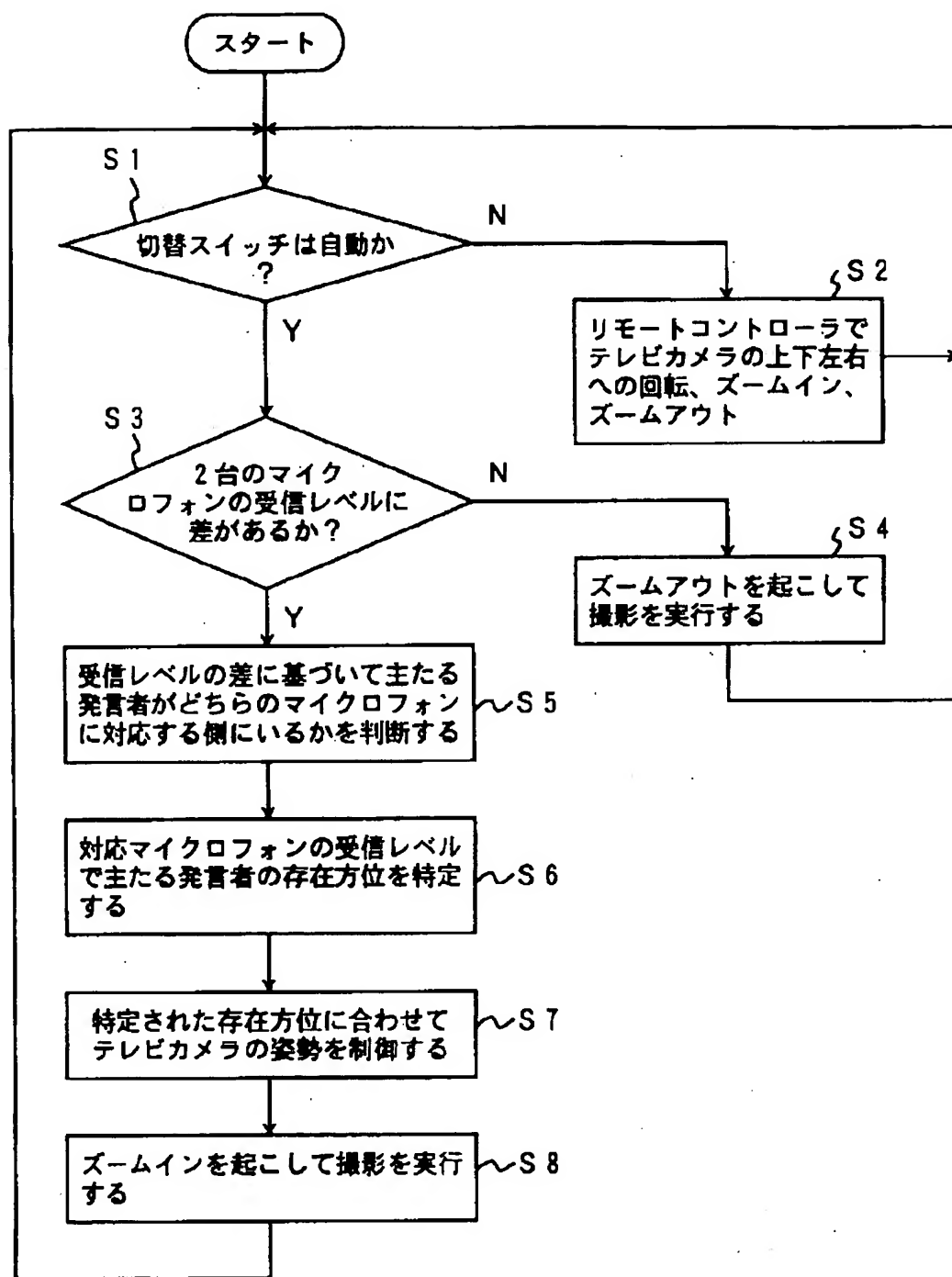


9: CRT switchy switch  
 6: Remote Controller  
 11: Display lamp

[Drawing 3]



[Drawing 2]



[Translation done.]